

Result	No.	Score	Query Match	Length	IDB	Description	Pred. No.

```

FEATURES
source      Biology, Pasadena, CA 91101, USA
            Location/Qualifiers
            1..2399
            /organism="Rattus norvegicus"
            /strain="Sprague Dawley"
            /clone="E13.5"
            279..1070
            /gene="DRG11"
            /codon_start=1
            /product="paired-like homeodomain transcription factor"
            /db_xref="PID:g1144015"
            /translation="MFTFPCPPQLEAGTAFGNHSTGDFDLRRKORNRRTFALQO
            LEAEVFAOYTHPDVTFRELANIKLNEARVQVWFNRRAKWRTERGASQDEPGA
            KEPAEYTPPPVNSPPGQDQARKEALEAOQSLGRVGPAGPFPSCLPCTLLN
            TATVAQALSHVLSKGLPCSCVDPDMGLSFLPYTCQSNRTASVAALRKAREHSE
            AVLASANLLPSTSSSPSPASKQVPPESQDKPSTKEQSEKSV"
            - polyA_site
            2399
            /note="25 A nucleotides"
BASE COUNT 511 a 681 c 658 g 549 t
ORIGIN
Jury Match          99.0%; Score 2399; DB 88; Length 2399;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 2399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1  gcagaggtaggcagggttcccgagcgctctcccggtccctgctctgggcttgggct 60
Qy 1  GCAGAGGTAGGAGGTTCCCGAGCGCTCTCCCGCTCCCTGCTCTGGGGCTTGGGGCT 60
Db 61  ccacggcttcttggcccgagctgctgctgctgagatggccttgcgcatgcgcggacc 120
Qy 61  CCACGGCTTCTTGGCCCGAGCTGCTGGCGCTGCAGATGGCTTGGCGGATCGCGGACC 120
Db 121  ccgctcggtgccaagtgcaggcttgcctggagacccctggaaccctggaaccagagcagac 180
Qy 121  CCGCTCGGTGCGCAAGTGCAGGCTTGTGGCTGGGACCCCTGAGAACCCAGAGGCCAGAC 180
Db 181  tgtgctcagcttgcagggcgagccagcagcgagcacaagctctgtagggcgccatcag 240
Qy 181  TGTGCTCAGCTTGCAGGCGGAGCCAGCCAGCGACAGGTCAGTCTGTGAGCGCCCATCAG 240
Db 241  tcttggtccagcgtcagggcccatccgacccgtcgagctggtttatttcccaactgcgcgc 300
Qy 241  TCTTGGTCTCAGCGCTCAGGCGCCATCCGACCGCTCGGGGATGTTTATTTCACCTGCCCGC 300
Db 301  cacagctagagggcacagcgttcttgaaccactcaggggatttgcagtagatgggt 360
Qy 301  CACAGCTAGAGGGCACAGCGCTTTTGGTAACCACTCTACGGGGGATTTTGTATGATGGGT 360
Db 361  ttcttagaagaaacagcgagaaatcggaacacttcgctcttcagcaggttgaagctc 420
Qy 361  TTCCTTAGAAGAAACAGCGCAGAAATCGGACAACTTCGCTCTTCACGATTTGGAAGCTC 420
Db 421  tggagcagttcttggcccaaacacactaccagatgtcttcacagagaagactagcca 480
Qy 421  TGGAGCAGTCTTTGGCCCAACACACTACCCAGATGTCTTCACAGAGAAGAGCTAGCCA 480
Db 481  tgaataaaacctcacagaagccagagtcaggtttggttccagaaaccgagagcaagt 540
Qy 481  TGAATAAAACCTCACAGAAGCCAGAGTGCAGGTTTGGTTCCAGAACCCGAGAGCCAAAGT 540
Db 541  ggaagaagacagagaggggctctgacacagaaaccaggggcttaagaaaccctagcag 600
Qy 541  GGAGGAAGACAGAGAGGGGCTCTGACCAAGNACCAGGGGGCTAAGGAACCCATGGCAG 600
Db 601  aggtgacacacccccagtggagaacatcaactctccacccccaggggaccagggccggg 660
Qy 601  AGGTGACACACCCCCAGTGAGGAACATCAACTCTCCACCCCCAGGGGACCCAGGCCGGG 660
Db 661  qcaagaagaggcccttggagccagcagagcctggagcacaagtggtggcccgccgggc 720
Qy 661  GCAAGAAGGAGGGCCCTGGAGGCCGAGCAGAGCCTGGGACGACAGTGGGGCCCGCGGGC 720

```

```

Db 721  ctttctccctcctctgttgcaggagccctctgaacacagcacttatgcccaggccc 780
Qy 721  CTTTCTTCCCTCTCTGTTGCCAGGACCCTCTCTAACACAGCCACTTATGCCAGGCC 780
Db 781  tgtcccatgtggcatctctgaagggggggcccaactgtgctcttctgctgctccagacccta 840
Qy 781  TGTCCCATGTGGCATCTCTGAAGGGGGGCCCACTGTGCTCTTGTGCTGCCAGACCTTA 840
Db 841  tggggctctctctctcccaacttacggttgcagagtaaacccagcagcagcgtggctg 900
Qy 841  TGGGGCTCTCTCTCTCCCACTTACGTTTGCAGAGTAACCCGACAGCCAGCGTGGCTG 900
Db 901  ccttgcgcatgaagggcccgagcattcagaagcgtctcgaagctgcacaccttctgc 960
Qy 901  CCTTGCGCATGAAGGGCCCGGAGCATTCAGAAGCGGTCTGTCAGTGTGCCAACCTTCTGC 960
Db 961  cgtccacagcagcagccccggccctgcctcccaagcaggtgcctccagagggcagcagg 1020
Qy 961  CGTCCACAGCAGCAGCCCGCGCCCTGCCTCCAAAGCAGGTGCTCCAGAGGCGAGCCAGG 1020
Db 1021  acaagcctcccaacgaaggaacagagcagggagagagagcgtatgaggtccggag 1080
Qy 1021  ACAAGCCCTCCCAACGAAGAAACAGAGCGAGGAGAGAGAGCGTATGAGGTCCGGAG 1080
Db 1081  aaccagctggggagccctgccaccctctctctcagcctcagccctgcagcctctg 1140
Qy 1081  AACCCAGCTGGGAGCCCTGCCACCCCTGCTCTCTCAGCCCTCAGCCCTGCCAGCCTCTG 1140
Db 1141  aacacaagagtagccacccctctcatgcatcgacggcagggcaaacgggacctgcaagctg 1200
Qy 1141  AACCAACAGAGTAGCCACCTCTCATGATCTGACAGGGCAACCGGGACCTGCACAGCTG 1200
Db 1201  gttgagacctgaagagtcctcctctagaaattctgctgtaggctgtgttctctcgtcttc 1260
Qy 1201  GTTGAGACCTGAAGAGTCCCTCTAGAATTCTGCTGTAGGCTGTGTGTCTCTCGCTTTC 1260
Db 1261  ctttgtgacatttccgtagctcttagtactctgacactgctctgtatgagagctc 1320
Qy 1261  CTTTGTGACATTTTCCGATGGCTCTTAGTACTCTGGACACTGCTCTGTGATGAGGTCC 1320
Db 1321  ctgtttttgctttttgttctctctcttttttttttttttttttttttttttttttttt 1380
Qy 1321  CTGTTTTTGTCTTTTGTCTCTCTTTTGTCTTTTGTCTTTTGTCTTTTGTCTTTTCCAGG 1380
Db 1381  caaagcagccttggaacaaagcagatttagttattccacactcctctcttgagatactctg 1440
Qy 1381  CCAAGCAGCCTTGGAGCAAAAGCAGATTAGTTTATTCACCATCTCTTCTTGAGATATCTG 1440
Db 1441  gaaggtcttgtcaattcccaagcactgtggcaaggatcatccctgaaagatgccaagaat 1500
Qy 1441  GAAGGTCTTGTCAATTCCCAAGGACTGTGGCAAGGATCATCCGTGAAAGATGCCAAGAT 1500
Db 1501  gacatcatgacagaaatgagacgggcactcccatattgcttaagaaccacagaaactg 1560
Qy 1501  GACATCTCATGACAGAAATGAGCGGCACTCCCATCTTCTTAAAGAACCAACAGAACTG 1560
Db 1561  gtgactatacagcaggttctcactcctcctcagccaggactggcctcctcttctgtct 1620
Qy 1561  GTGGACTATCAGCCAGTCTCCTACCTCTCAGCCAGGACTGGCCTCGGCTCCTCTTGTCT 1620
Db 1621  tgtttaaagaaattagctgaggttttggctagaaagtgcacagatatgggctgagagact 1680
Qy 1621  TGTTTAAAGAAATTAGCTGAGGTTTGGCTAGGAAAGTGAAGATATGGGCTGAAGACAT 1680
Db 1681  tgtgtctgacctagcagatctccctggggcacatctgacctgtgcaagtcaggcaggt 1740
Qy 1681  TGTGTCTGACCTTAGCAGATCTCCCTGGGGCACATCTGACCTGGTCCAGTCAGGCAGGT 1740
Db 1741  tgtcagttcggggatgaggggctgtctctgtctgtatctgtgtgtgtgttccctcgaataga 1800
Qy 1741  TGTCAAGTTCGGGATGGGGCTGCTCTGTGTATCTGTGTGTGGGTCTCCTCCTCAATTAGA 1800
Db 1801  gttgttaccttgagggcccgctctcttcagaagagtgatgggaaagtccaccatcagaat 1860

```

Query Match 3.4%; Score 82; DB 78; Length 1433;
 Best Local Similarity 71.6%; Pred. No. 1.07e-22;
 Matches 136; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

```

Ddb 220 agaagcagcgaagaataagacaaaccttcaatagcagccagctgcagctttagagcgt 279
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
369 AGAAACAGCGCAGAAATCGGCAACCTTCGTCTCTTCAGCAGCTTGGAGCTCTGGAGGCA 428
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
280 gtctttgagcgacacactactctctgctcttttgcgagagacacctgcccgcgggtg 339
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
429 GTCCTGCCAACACACTACCAGATGCTTCACAGAGAGAGACTAGCATGAATA 488
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
340 aacctaccgagcgagagtcgaggtgtgtgttcagacccgaagcgaagtccgcagg 399
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
489 AACCTACAGAAAGCCAGAGTGCAGGTTTGTTCCAGAACCGAAGACCAAGTGGAGGA 548
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Ddb 400 aatgagagag 409
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 549 ACAGAGAGAG 558
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  
```

3 MUSCHX10A 3089 bp mRNA ROD 12-JAN-1995
 LOCUS Mus musculus (clone 6) homeobox protein (Chx10) mRNA, complete cds.
 DEFINITION L34808
 ACCESSION L34808
 KEYWORDS homeobox protein.
 SOURCE Mus musculus (strain CD1/129sv) cDNA to mRNA.
 ORGANISM Mus musculus
 Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
 Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
 Murinae; Mus.
 1 (bases 1 to 3089)
 Liu, I. S., Chen, J. D., Ploder, L., Vidgen, D., van der Kooy, D.,
 Kamin, Y. I., and McInnes, R. R.
 Developmental expression of a novel murine homeobox gene (Chx10):
 evidence for roles in determination of the neuroretina and inner
 nuclear layer
 Neuron 13 (2), 377-393 (1994)

JOURNAL 54338594
 MEDLINE
 FEATURES Location/Qualifiers
 source
 1..3089
 /organism="Mus musculus"
 /strain="CD1/129sv"
 /map="chromosome 12"
 1..1086
 /gene="Chx10"
 /codon_start=1
 /product="homeobox protein"
 /db_xref="PID:g623166"
 /translation="MTGKAGALSFKPKSTVAKSTSGGAPARCTGFGIQEILGLINKP
 PSSHPRAADGLAPCHLLAARSVLSPAVGSNGLLGPGGLPGFYTQPTFEVLSDPS
 VHLPQPGAGPDLDTSTQASSDSEVSSDRKMSKALNQPKKKRRHRTFTSYQL
 EELEKAFNAFNPVDVVTARELANKTELPEDRIQVFNRRKRRKRCWGRSSVMAE
 YGLYAWNPISPLPSILSKARDGIMDSQAPMLGLMHKSKLEAAASGRKRPEVERQA
 LPLDKRMEBEREAPAAQAIISOELRENSIALRAKAQEHSTKVLGTVSGPDSILARN
 EKPEEEDATEEDRPAEKLSPQLEDNA"

polyA_site 3089
 BASE COUNT 703 a 897 c 800 g 689 t
 ORIGIN

Query Match 3.3%; Score 80; DB 84; Length 3089;
 Best Local Similarity 72.2%; Pred. No. 1.02e-21;
 Matches 130; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

```

Ddb 443 agcgcgacacaggaacaatcttacttccaccagctagagagctggagaagcattca 502
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 376 ACCGCAGAAATCGGACAACCTTCCTCTTCAGCAGTTGGAAGCTCTGGAGGAGCTTTG 435
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Ddb 503 atgaagccactaaccagatgtctacgcccgggagatgctgcccataagaaaggagctcc 562
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  
```


CDS

Query Match 3.0%; Score 72; DB 86; Length 1760;
 Best Local Similarity 69.1%; Pred. No. 7.40e-18;
 Matches 130; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Db 845 agcagcaaaagagagagacacacactcgttccaccagtttgacgtgtagagaaactagag 904
 QY 366 AGAAGAAAACAGCGCAGAGATCGGACAACTTCGCTCTTCAGCAGCTTGAAGCTCTGGAG 425
 Db 905 aaggtcttccaaaacacattaccgcgatgtatgtcagagacacagcttgcaactgaga 964
 QY 426 CGAGTCTTTGGCCAAACACACTACCAGATGCTTTCACAGAGAGAGCTAGCCATGAAA 485
 Db 965 acggagctcactgagcgagggtccaggttgggttcaaaatcgaaggcccaagtggaga 1024
 QY 486 ATAACTTCACAGAGCAGAGTGCAGGTGGTTGTTCCAGAACCGAGAGCCCAAGTGGAGG 545
 Db 1025 aaagagaga 1032
 QY 546 AAGACAGA 553

RESULT 9
 LOCUS MMU03873 3359 bp mRNA ROD 16-MAR-1994
 DEFINITION Mus musculus homeobox (Pmx) mRNA, complete cds.
 ACCESSION U03873
 NID 9460124
 KEYWORDS house mouse.
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 3359)
 AUTHORS Kern, M.J., Argao, E.A., Birkenmeier, E.H., Rowe, L.B. and Potter, S.S.
 TITLE Genomic organization and chromosome localization of the murine homeobox gene Pmx
 JOURNAL Genomics 19 (2), 334-340 (1994)
 MEDLINE 94245205
 REFERENCE 2 (bases 1 to 3359)
 AUTHORS Kern, M.J.
 TITLE Direct Submission
 JOURNAL Submitted (01-DEC-1993) Michael J Kern, Basic Science Research, Childrens Hospital, Research Foundation, 3333 Burnet Ave, Cincinnati, OH 45229-3039, USA

FEATURES
 Location/Qualifiers
 source 1..3359
 /organism="Mus musculus"
 5'UTR 1..1021
 CDS 1022..1675
 /gene="Pmx"
 /codon_start=1
 /product="homeobox protein"
 /db_xref="PID:9460125"
 /translation="MTSSYGHVLERQPALGRRLDSPGNLPTLQAKNFVSVSHLLDLEE AGWAAQADESVGEAGRLSPGLTSGSDTPQONDQLNSEKKRKRORNRRTFN SSOALQERFERTHYDAFVREDLARVNLTEAROVQVFNRRFRNERNRLANK NASLLKSYSGDVTAVEQPIVPRAPRTDYLWGTSAPYSSSLPRCCLHEGLHNGF" order(1262,1438,1620,1692)
 /note="introns appear at these places in the genomic sequence"
 3'UTR 1676..3359
 polyA_signal 1844
 polyA_signal 1925
 polyA_signal 2158
 polyA_signal 2360
 polyA_signal 2486
 stem_loop 2665..2717
 /note="stem 19/20 bp and loop 13 bp"
 polyA_signal 2916
 polyA_signal 3111
 BASE COUNT 891 a 771 c 736 g 961 t
 ORIGIN

Query Match 3.0%; Score 72; DB 82; Length 3359;
 Best Local Similarity 68.9%; Pred. No. 7.40e-18;
 Matches 131; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

Db 1295 agaaagcagcgagagagacacacattcaacagcagcagcagcagcagcagcagcagcgt 1354
 QY 369 AGAAAACAGCGCAGAAATCGGACAACTTCGCTCTTCAGCAGTTGGAAAGCTCTGGAGGCA 428
 Db 1355 gtotttgagcgagacacattaccgcgatgttttgcgttcgagagagatctcgcaactcggggtg 1414
 QY 429 GTCTTTGCCCAACACACTACCAGATGCTTTCACAGAGAGAGCTAGCCATGAAAAATA 488
 Db 1415 aactcactgagcgagagtcaggtggtgttttcagacacagcagcagcagcagcagcagcag 1474
 QY 489 AACTCAGAGAGCCAGAGTGCAGGTGGTTGTTCCAGAACCGAGAGCCCAAGTGGAGGAG 548
 Db 1475 aatgagcgag 1484
 QY 549 ACAGAGAGAG 558

RESULT 10
 LOCUS CAU07056 1588 bp mRNA VRT 22-OCT-1994
 DEFINITION Carassius auratus homeobox protein mRNA, complete cds.
 ACCESSION U07056
 NID 9460927
 KEYWORDS goldfish.
 SOURCE Carassius auratus
 ORGANISM Carassius auratus
 Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes; Cyprinidae; Cyprininae; Carassius.
 REFERENCE 1 (bases 1 to 1588)
 AUTHORS Levine, E.M., Hitchcock, P.F., Glasgow, E. and Schechter, N.
 TITLE Restricted expression of a new paired-class homeobox gene in normal J. Comp. Neurol. 348 (4), 596-606 (1994)
 JOURNAL 95138312
 MEDLINE
 REFERENCE 2 (bases 1 to 1588)
 AUTHORS Levine, E.
 TITLE Direct Submission
 JOURNAL Submitted (23-FEB-1994) Edward M. Levine, Health Sciences Center, SUNY at Stony Brook, Stony Brook, NY 11794, USA

FEATURES
 Location/Qualifiers
 source 1..1588
 /organism="Carassius auratus"
 /tissue_type="retina"
 /dev_stage="adult"
 CDS 30..1055
 /codon_start=1
 /product="homeobox protein"
 /db_xref="PID:9460928"

Query Match 2.9%; Score 71; DB 49; Length 1588;
 Best Local Similarity 69.0%; Pred. No. 2.22e-17;
 Matches 129; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Db 465 agaaaaagcagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 524
 QY 369 AGAAAACAGCGCAGAAATCGGACAACTTCGCTCTTCAGCAGTTGGAAAGCTCTGGAGGCA 428
 Db 525 gccttcacgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 584
 QY 525 gccttcacgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 584

```

QY 429 GTCTTGGCCCAACACACTACCCAGATGCTTCCACAGAGAGAGAGTACCCATGAAATA 488
Db 585 gagctccagagacagaatacaggtatgtttcagaacccgctgctgaaggagaa 644
QY 489 AACCTCAGAGAGCCAGAGTGCAGGTTTGGTTCCAGAACCGAGAGCCAGTGGAGGA 548
Db 645 agggaga 651
QY 549 ACAGAGA 555

RESULT 11
LOCUS MALLX3 1744 bp RNA 30-MAR-1995
DEFINITION M.auratus mRNA for alx3.
ACCESSION X81403
NID g587455
KEYWORDS alx3 gene.
SOURCE golden hamster.
ORGANISM Mesocricetus auratus
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
Cricetinae; Mesocricetus.
REFERENCE 1 (bases 1 to 1744)
AUTHORS German,M.S.
TITLE Direct Submission
JOURNAL
FEATURES
Source
1..1744
/organism="Mesocricetus auratus"
/tissue_type="pancreatic islet"
/cell_type="beta-cell"
/cell_line="Hir-T15.M2.2.2"
/clone_lib="lambda gt11"
/clone="MG 1343.1"
21..1061
/gene="alx3"
/db_xref="PID:g587456"
/db_xref="PID:g587456"
/translation="MAPDPERCAPFVGPAGPVAAGDPPGPGQSDASPHLHPA
PPRPLSRFPACGLEPLPEPAKPKYLQDLGPAPVINGHFYEGPAEAEKASK
AASFQPLVDCRGGRDPSNVGSGFGLASLRVPLSPGLPDMELAKSKRRNR
TTFSTQLEELKVEFKTHYPDVYAREQLALTDTEARVQVFNRRARWRKRYG
KMQEGRFTTAYDIFSLPRDSDHPQLNSLWPGSGSGPGCLMSPEGIPSCMSP
YSHGNGVAGFMGVPAAPAHGPIYSIHGFPALGHSFEPSPDGYKSPSLSVRYK
PREPPLNLNWT"
BASE COUNT 365 a 559 c 472 g 348 t
ORIGIN
Query Match 2.9%; Score 70; DB 81; Length 1744;
Best Local Similarity 69.4%; Pred. No. 6.65e-17;
Matches 125; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Db 480 agcaagaagcgtcggaaccgcacacattccagtcggaagagctggagaa 539
QY 369 AGAAAACAGCCGAGAAATCGACAACCTTCGCTCTTCAGCAGTGTGAAGCTCTGGAGCA 428
Db 540 gtcttcagaaaccactaccgcagctgtagtcccgagagcagctggttaagcag 599
QY 429 GTCTTGGCCCAACACACTACCCAGATGCTTCCACAGAGAGAGTACCCATGAAATA 488
Db 600 gacctgacagagcccggttacagctgtgtttccagaacccgagaccagtgccggaag 659
QY 489 AACCTCAGAGAGCCAGAGTGCAGGTTTGGTTCCAGAACCGAGAGCCAGTGGAGGA 548

RESULT 12

```

```

LOCUS GPRX2 849 bp RNA 29-MAR-1996
DEFINITION G.gallus Prx-2 (S8) mRNA.
ACCESSION X79695
NID g558376
KEYWORDS Prx-2 gene.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Archosauria; Aves; Neognathae; Galliformes;
Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 849)
AUTHORS Leussink,B., Brouwer,A., el Khattabi,M., Poelmann,R.E.,
Gittenberger-de Groot,A.C. and Meijlink,F.
TITLE Expression patterns of the paired-related homeobox genes MHox/Prx1
and S8/Prx2 suggest roles in development of the heart and the
forebrain
JOURNAL Mech. Dev. 52 (1), 51-64 (1995)
MEDLINE 96076127
REFERENCE 2 (bases 1 to 849)
AUTHORS Meijlink,F.
TITLE Direct Submission
JOURNAL Submitted (10-JUN-1994) F. Meijlink, Netherlands Inst. for
Developmental, Biology, Hubrecht Laboratory, Uppsalalaan 8, 3584 CT
Utrecht, NETHERLANDS
FEATURES
Source
1..849
/organism="Gallus gallus"
/dev_stage="10 days"
/clone_lib="lambda gt11 chicken embryo libraries one from
Clontech and one from Karolinska Institute"
<1..498
/gene="prx-2 (S8)"
/codon_start=1
/db_xref="PID:g558377"
/translation="ESLSPRGVAKRKKKORRNTTENSQQLALERVFERHTHPDAF
VREELARVNLSEARVQVFNRRARERMLANRSLKSYSOEAAIEQPMAP
RPTALSPYLSSSSSPYSTVPSYSSGTATAGGVNMANSLRLKAKEFSLHQNO
VPTVN"
BASE COUNT 208 a 266 c 222 g 153 t
ORIGIN
Query Match 2.8%; Score 69; DB 50; Length 849;
Best Local Similarity 70.4%; Pred. No. 1.98e-16;
Matches 119; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Db 43 aaacagagagaaacccacacattccacagcagccacagctcagagaggttc 102
QY 372 AACAGCGCAGAAATCGGACACCTTCCTCTTCCAGCAGTTGGAAGCTCTGGAGCGATC 431
Db 103 tttagcgggacacattaccccgagccttcgtcgggagagagctgcccagagaggtgaac 162
QY 432 TTTGCCCAACACACTACCCAGATGCTTCCACAGAGAGAGTACCCATGAAATAAC 491
Db 163 ctacgagcggcagagtcaggtcttggtttcagaacccaggggccaagt 211
QY 492 CTCACAGAAGCCAGAGTGCAGGTTTGGTTCCAGAACCCGAGAGCCCAAGT 540

RESULT 13
LOCUS gMHox-homeobox 895 bp mRNA 22-SEP-1994
DEFINITION gMHox-homeobox (chickens, stage 17 embryo, heart, mRNA, 895 nt).
ACCESSION S69088
NID g545359
KEYWORDS chickens heart stage 17 embryo.
SOURCE Gallus sp.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 895)
AUTHORS Kuratani,S., Martin,J.F., Wawersik,S., Lilly,B., Eichele,G. and
Olson,E.N.
TITLE The expression pattern of the chick homeobox gene gMHox suggests a
role in patterning of the limbs and face and in
compartmentalization of somites

```

JOURNAL Dev. Biol. 161 (2), 357-369 (1994)
MEDLINE 94148118
REMARK GenBank staff at the National Library of Medicine created this entry [NCBI gi533 144347] from the original journal article.
This sequence comes from Fig. 1A.

FEATURES
source Location/Qualifiers
1..895 /organism="Gallus sp."
CD5 31..768 /gene="gmHox"
/note="Description: homeobox; homeodomain protein; Method:
conceptual translation supplied by author. This sequence
comes from Fig. 1A"
/codon_start=1
/db_xref="pid:g545360"
/translation="NASSVAHAMERQALLPARLDGPAGLNDLOAKNFVSVSHLLDLEE
AGDMAACDEGGSPGLSLPGLTSGSDTPQOONDOLNSEKKKPKORNTFTFN
SSQLQALRVERFTHYPDAFYREDIARRVNLTARVQVFQNRRAKFRNERAMLAASK
NSLJKSTGDVTAEVEYIPAPRPPTDIYLSWGTASPYSAMATYSTTTCTNASPAQGG
NNANIANLRKAKIKSYQRNPVPTN"

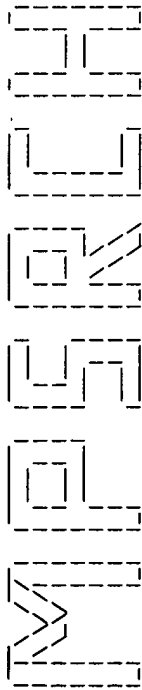
BASE COUNT 254 a 255 c 246 g 140 t
/ORIGIN

Query Match 2.98; Score 67; DB 51; Length 895;
Best Local Similarity 69.4%; Pred. No. 1.75e-15;
Matches 120; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

Db 303 aaaggaagcagagaagaaacaggactacccttcaaacagcagccagctccaggcactagagag 362
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 368 AGAAACAGCGCAGGAANYCGGACACCTTCGCCTCTTCAGCACTTGAGAGCTCTGGAGGC 427
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 363 ggctcttgagagagacacataccccgatgcctttgtcacgggaagacctgacgcagagt 422
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 428 AGTCTTTGCCAAAACACACTACCACCATGTCTTCACAGAGAAGAGCTAGCCATGAAAT 487
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 423 caacctactgaagccagaggttcaggtgtggttcacgaaccggagggccaagt 475
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 488 AAACCTCAGAGAAGCCAGAGTGCGAGGTTGGTTCCAGAACCCGAGAGGCCAAGT 540
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
LOCUS CHKPRX1 1084 bp mRNA VRT 10-SEP-1993
DEFINITION Gallus gallus paired-related homeobox mRNA, complete cds.
ACCESSION D13433
ENTRY 9222850
KEYWORDS paired-related homeobox.
SOURCE Gallus gallus (library: lambda gt10) stage 24-26 CDNA to mRNA,
clones p2 and p7.
ORGANISM Gallus gallus
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Archosauria; Aves; Neognathae; Galliformes;
Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1084)
AUTHORS Nohno,T., Koyama,E., Myokai,F., Taniguchi,S., Ohuchi,H., Saito,T.
and Noji,S.
TITLE The chicken homeobox gene related to Drosophila paired is
predominantly expressed in the developing limb
JOURNAL Dev. Biol. 158, 254-264 (1993)
MEDLINE 9331789
COMMENT Submitted (20-OCT-1992) to DBJ by: Tsutomu Nohno
Department of Pharmacology
Kawasaki Medical School
577 Matsushima
Kurashiki 701-01
Japan
Phone: 086-462-1111
Fax: 086-462-1199.
Location/Qualifiers
1..1084
/organism="Gallus gallus"
CD5 233..967
/gene="prx-1"
/codon_start=1

/product="paired-related homeotic gene product"	
/db_xref="PID:G222851"	
/translation="MASSAYAMEROALLPARLDQAGLDNLQAKNFSVSHLLDLEE	
/translation="AGDVAAGDDEGGEGPGRSLLESPLTSGSTPOODNDOLNSEKKKQKORNRRTFN	
SSQALRVRERHYHPDAFYREDLARVNLTEARVQFVQNRRAKFRNERNLASKN	
ASLKSUSGSDYAVEQPIVPRAPRPDTILSWGTSASPYSAMATYSTCTTNASPAQGMN	
MANSTIANLRUKAKESYLGQRNVPTVN"	
misc_feature	
509..691	
/gene="Prx-1"	
/note="homeobox"	
BASE COUNT 258 a 337 c 317 g . 172 t	
ORIGIN	
Query Match 2.8%; Score 69; DB 49; Length 1084;	
Best Local Similarity 69.9%; Pred. No. 1.98e-16;	
Matches 121; Conservative 0; Mismatches 52; Indels 0; Gaps 0;	
Db	505 agggaagcagagaagaaacagagactctcaacagcagcagcagcagcagcagagag 564
QY	368 AAGAAACAGCGCAGAAATTCGGAACACCTTCCTCTTCAGCAGCTTGGAAAGCTCTGGAGGC 427
Db	565 ggtctttgagaggacacactaccgcatccttggtagcgggaagacctgtgcagcagagt 624
QY	428 AGTCCTTGCCCAACACACTACCAGATGCTTCACCAGAGAGAGCTAGCCATGAAAT 487
Db	625 caacctcactgaagccagagtgtagtggtggtccagaacccgagagccaaagt 677
QY	488 AAACCTCAGAGAAGCCAGAGTCAGGTGGTTCAGCAACCCGAGAGCCCAAGT 540
RESULT 15	XLXOTX2S1 1264 bp mRNA VRT 09-FEB-1995
LOCUS	Xenopus laevis homeobox protein orthodenticle 2 (Xotx2) mRNA,
DEFINITION	complete cds.
ACCESSION	U19813
NID	9644779
KEYWORDS	1 of 2
SEGMENT	African clawed frog.
SOURCE	Xenopus laevis
ORGANISM	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
	Vertebrata; Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea;
	Pipidae; Xenopodinae; Xenopus.
REFERENCE	Blitz,I.L. and Cho,K.W.Y.
AUTHORS	Anterior neuroectoderm is progressively induced during gastrulation:
TITLE	the role of the Xenopus homeobox gene orthodenticle
JOURNAL	Development (1995) in press
REFERENCE	2 (bases 1 to 1264)
AUTHORS	Blitz,I.L.
TITLE	Direct Submission
JOURNAL	Submitted (12-JAN-1995) Ira Blitz, Developmental and Cell Biology
	and the Developmental Biology Center, University of California,
	Irvine, CA 92717-2300, USA
FEATURES	Location/Qualifiers
source	1..1264
	/organism="Xenopus laevis"
5'UTR	<1..212
CDS	213..1082
	/gene="Xotx2"
	/note="misexpression of Xotx2 induces ectopic cement
	glands; homeodomain protein"
	/codon_start=1
	/product="orthodenticle 2"
	/db_xref="PID:9644782"
	/translation="MMSLKQPPYAVNGLSLTTSGMDLLHPVSGVPATPRKORRERTT
	FTRAQDLLEALFAKTPDIFMEAEVAKINLPESRVQVWFKNRAKCRQOQOQOQ
	GQKNRVRSKKTPSPAREVSESTSGQSPFISSTGPVISTAPVTSWSPASISPL
	SDPLSTNATVSHLQSGYPTWTQASGYSQVASTSYFGMDGCSYLTPMHOLSGPGAT
	LPKSTNATVSHLQSGYPTWTQASGYSQVASTSYFGMDGCSYLTQDTSKWLNFNACDL
	IKDQTSWKNQVFL"
	1083..>1264
3'UTR	335 a 366 c 278 g 285 t
BASE COUNT	



(TM)

Release 2.1d John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by Intelligenetics, Inc.

MPSrch_n n.a. - n.a. database search, using Smith-Waterman algorithm

'in on: Fri May 30 16:06:24 1997; MasPar time 1700.26 Seconds
912.837 Million cell updates/sec
-dbular output not generated.

Title: >US-08-701-278-1
Description: (1-2424) from US08701278.seq
Perfect Score: 2424
N.A. Sequence: 1 GCAGAGGTAGGACGGGTTC.....GAAAGAAAAA.....CTTTTTTTTTTTTT
Comp: CGTCTCCATCGTCCACGAGG.....

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 886179 seqs, 320143548 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: EST-STS
1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86
87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:EST92
93:EST93 94:EST94 95:EST95 96:EST96 97:EST97 98:EST98
99:EST99
EST-STS-TWO
100:EST100 101:EST101 102:EST102 103:EST103 104:EST104
105:EST105 106:EST106 107:EST107 108:EST108 109:EST109
110:EST110 111:EST111 112:EST112 113:EST113 114:EST114
115:EST115 116:EST116 117:EST117 118:EST118 119:EST119
120:EST120 121:EST121 122:EST122 123:EST123 124:EST124
125:EST125 126:EST126 127:EST127 128:EST128 129:EST129
130:EST130 131:EST131 132:EST132 133:EST133 134:EST134
135:EST135 136:EST136 137:EST137 138:EST138 139:EST139
140:EST140 141:EST141 142:EST142 143:EST143 144:EST144
145:EST145 146:EST146 147:EST147 148:EST148 149:EST149
150:EST150 151:EST151 152:EST152 153:EST153 154:EST154
155:EST155 156:EST156 157:EST157 158:EST158 159:EST159
160:EST160 161:EST161 162:EST162 163:EST163 164:EST164
165:EST165 166:EST166 167:EST167 168:EST168 169:EST169
170:EST170 171:EST171 172:EST172 173:EST173 174:EST174

Database:

EST-STS-THREE

175:EST175 176:EST176 177:EST177 178:EST178 179:EST179
180:EST180 181:EST181 182:EST182 183:EST183 184:EST184
185:EST185 186:EST186 187:EST187 188:EST188 189:EST189
190:EST190 191:EST191 192:EST192 193:EST193
194:STS1 195:STS2 196:STS3 197:STS4 198:STS5 199:STS6
200:STS7 201:STS8 202:STS9 203:STS10 204:STS11 205:STS12
206:gnEST1 207:gnEST2 208:gnEST3 209:gnEST4 210:gnEST5
211:gnEST6 212:gnEST7 213:gnEST8 214:gnEST9 215:gnEST10
216:gnEST11 217:gnEST12 218:gnEST13 219:gnEST14
220:gnEST15 221:gnEST16 222:gnEST17 223:gnEST18 224:gnEST19
225:gnEST20 226:gnEST21 227:gnEST22 228:gnEST23 229:gnEST24
230:gnEST25 231:gnEST26 232:gnEST27 233:gnEST28 234:gnEST29
235:gnEST30 236:gnEST31 237:gnEST32 238:gnEST33 239:gnEST34
240:gnEST35 241:gnEST36 242:gnEST37 243:gnEST38 244:gnEST39
245:gnEST40 246:gnEST41 247:gnEST42 248:gnEST43 249:gnEST44
250:gnEST45

Statistics: Mean 12.474; Variance 4.448; scale 2.805

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description	Pred. No.
C 1	76	3.1	498	26	AA022576	ze72e10.r1 Soares fet 2.79e-46
C 2	71	2.9	829	219	W22594	70B5 Human retina CDN 3.45e-41
C 3	71	2.9	829	186	W22594	70B5 Human retina CDN 3.45e-41
C 4	61	2.5	324	125	R95996	yt83c02.r1 Homo sapie 3.10e-31
C 5	60	2.5	386	131	T09185	EST07078 Homo sapiens 2.94e-30
C 6	55	2.3	572	240	MAA33108	mi26g03.r1 Soares mou 1.92e-25
C 7	55	2.3	572	3	AA033108	mi26g03.r1 Soares mou 1.92e-25
C 8	55	2.3	604	182	W89306	mf81d10.r1 Soares mou 1.92e-25
C 9	54	2.2	424	185	W17990	mb82e05.r1 Soares mou 1.71e-24
C 10	54	2.2	424	217	W17990	mb82e05.r1 Soares mou 1.71e-24
C 11	53	2.2	513	26	AA022577	ze72e10.r1 Soares fet 1.51e-23
C 12	51	2.1	439	172	W59350	md49b07.r1 Soares mou 1.12e-21
C 13	49	2.0	372	128	SSC1E11	S.scrofa mRNA; expres 7.98e-20
C 14	49	2.0	372	249	SSC1E11	S.scrofa mRNA; expres 7.98e-20
C 15	45	1.9	309	180	W81935	me93d11.r1 Soares mou 3.39e-16
C 16	47	1.9	488	97	N99100	zb81a08.r1 Homo sapie 5.36e-18
C 17	47	1.9	488	211	N99100	zb81a08.r1 Homo sapie 5.36e-18
C 18	44	1.8	303	13	AA059929	mf75d03.r1 Soares mou 2.63e-15
C 19	44	1.8	303	242	MAA59929	mf75d03.r1 Soares mou 2.63e-15
C 20	44	1.8	323	185	W18334	mb68b03.r1 Soares mou 2.63e-15
C 21	44	1.8	323	218	W18334	mb68b03.r1 Soares mou 2.63e-15
C 22	43	1.8	448	220	W24124	zb81a08.r1 Soares sen 2.00e-14
C 23	43	1.8	448	186	W24124	zb81a08.r1 Soares sen 2.00e-14
C 24	42	1.7	154	240	MAA32943	mi26g08.r1 Soares mou 1.50e-13
C 25	42	1.7	154	3	AA032943	mi26g08.r1 Soares mou 1.50e-13
C 26	40	1.7	172	178	W77302	me65f05.r1 Soares mou 1.73e-08
C 27	36	1.5	78	185	W19236	yb90c01.r1 Homo sapie 1.73e-08
C 28	36	1.5	255	36	H28635	y152d08.r1 Homo sapie 1.73e-08
C 29	37	1.5	260	43	C16221	Human aorta CDNA 5'e 2.63e-09
C 30	37	1.5	260	237	HSC2219	Human aorta CDNA 5'e 2.63e-09
C 31	37	1.5	273	179	H80149	me89g06.r1 Soares mou 2.63e-09
C 32	37	1.5	274	40	H38014	YP58g02.r1 Homo sapie 2.63e-09
C 33	36	1.5	306	96	N95566	YP60f09.r1 Homo sapie 1.73e-08
C 34	36	1.5	327	154	W34141	mb81f03.r1 Soares mou 1.73e-08
C 35	37	1.5	332	127	R1CC0082B	Rice CDNA, partial se 2.63e-09
C 36	37	1.5	335	16	AA067108	mm32c09.r1 Stratagene 2.63e-09
C 37	37	1.5	335	244	MAA67108	mm32c09.r1 Stratagene 2.63e-09
C 38	36	1.5	337	196	G07625	human STS SHGC-5727 c 1.73e-08
C 39	36	1.5	392	42	C14283	Human fetal brain CDN 1.73e-08
C 40	36	1.5	392	237	HSC2837	Human fetal brain CDN 1.73e-08
C 41	36	1.5	400	168	W48910	mc20f11.r1 Life Tech 1.73e-08
C 42	36	1.5	428	43	C14427	Human fetal brain CDN 1.73e-08
C 43	37	1.5	430	237	HSC3947	Human fetal brain CDN 2.63e-09

	Query Match	2.98	Score 71	DB 219	Length 829
	Best Local Similarity	70.48	Pred. No. 3.45e-41		
	Matches 119	Conservative 0	Mismatches 50	Indels 0	Gaps 0
Db	116	acttagctcttcgattctttaaacacatactcgacctcgactcgccgacngttagttttca	175		
Cp	540	ACTTGGCTCTTCGGTTCTTGGGAACCAACCTGCACCTCGGCTTCTGTGAGCTTTATTTC	481		
Db	176	gtgccacctctctcgcatgaagatctctgggtggtaccgggtctctggcaaacagtgcttcca	235		
Cp	480	TGGCTAGCTTCTCTGTTGGTAACACATCTGGGTAGTGTGTTGGGCAAGACTGCCTCCA	421		
Db	236	gcacatctagctgcgcgcgagtgaaacgtctctctccccgngactgttt	284		

CP 420 GAGCTCCAACTGCTGAAGAGCGAAGGTTGTCGAGTTCTTCGCGCTGTTT 372

RESULT 3
LOCUS W22594 829 bp mRNA EST 06-MAY-1996
DEFINITION 7085 Human retina cDNA Tsp5091-cleaved sublibrary Homo sapiens cDNA not directional.

ACCESSION W22594
NID 91399427

KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens

Eukaryotae; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 829)

REFERENCE Macke, J., Smallwood, P. and Nathans, J.

TITLE Adult Human Retina cDNA

JOURNAL Unpublished (1996)

COMMENT

Contact: Dr. Jeremy Nathans
Dr. Jeremy Nathans, Dept. of Molecular Biology and Genetics
Johns Hopkins School of Medicine
725 North Wolfe Street, Baltimore, MD 21205

Tel: 410 955 4678

Fax: 410 614 0827

Email: jeremy.nathans@gmail.bs.jhu.edu

Clones from this library are NOT available.

PCR Primers

FORWARD: CTTTGTGACAGTTTCAGCTGCTTAAGT

BACKWARD: GAGTGGCTTATGAGTATTTCTTCAGGGTAA

Seq primer: GGGTAAAGCAAGAAAT.

Location/Qualifiers

1..829

/organism="Homo sapiens"

/note="Organ: eye; Vector: lambda gt10; Site_1: EcoRI; Site_2: EcoRI; The library used for sequencing was a sublibrary derived from a human retina cDNA library. Inserts from retina cDNA library DNA were isolated, cleaved with Tsp5091, size selected, and cloned into lambda gt10. Individual plaques were arrayed and used as templates for PCR amplification and these PCR products were used for sequencing."

/clone.lib="Human retina cDNA Tsp5091-cleaved sublibrary"

/sex="mixed (males and females)"

/tissue_type="retina"

/dev_stage="adult"

/lab_host="E. coli strain K802"

<1..>829

mRNA 155 a 184 c 218 g 194 t 78 others

BASE COUNT

ORIGIN

Query Match 2.9%; Score 71; DB 186; Length 829;

Best Local Similarity 70.4%; Pred. No. 3.45e-41;

Matches 119; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Db 116 acttgctcttcgattcttaaacaccatccctgactcggcangttgatttca 175

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

CP 540 ACTTGCTCTTCGGTCTCGAACCAACCTGCACCTGCTCTCTGAGGTTATTTC 481

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 176 gtgcacacctctctcgcagatgaagatgtctgggtacccggtcttggcaacagtgctcca 235

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

CP 480 TGGCTAGTCTCTCTGTTGGAAGACATCTGGGTAGTGTGTTGGGCAAGACTGCTCCA 421

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 236 gacatctagctgcgcccgagtgacagctgtctctctccggngctgttt 284

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

CP 420 GAGCTCCAACTGCTGAAGAGCGAAGGTTGTCGAGTTCTTCGCGCTGTTT 372

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4

LOCUS R95996 324 bp mRNA EST 11-SEP-1995

DEFINITION Yt83c02.r1 Homo sapiens cDNA clone 230882 5' similar to gb:M93650 PAIRED BOX PROTEIN PAX-6 (HUMAN);

ACCESSION R95996

NID Yt83c02

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryotae; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 324)

REFERENCE Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.

TITLE The WashU-Merck EST Project

JOURNAL Unpublished (1995)

COMMENT

ACCESSION R95996

NID 9581656

KEYWORDS EST.

SOURCE

clone=230882 primer=W13RP1 library=Soares pineal gland N3HPG vector=PT73D (pharmacia) with a modified polylinker host=DH10B (ampicillin resistant) Rsite1=Not I Rsite2=Eco RI 1st strand CDNA was primed with a Not I - Oligo(df) primer [5' TTTTACCAATCTGAAGTGGAGCGCGCGCTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo.

ORGANISM Homo sapiens

Eukaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcoterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 324)

REFERENCE Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.

TITLE The WashU-Merck EST Project

JOURNAL Unpublished (1995)

COMMENT

Contact: Willson RK

WashU-Merck EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: estwatson.wustl.edu

High quality sequence stops: 180

Source: IMAGE Consortium, LBNL

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Location/Qualifiers

1..324

/organism="Homo sapiens"

/clone="230882"

<1..>324

mRNA 104 a 74 c 71 g 61 t 14 others

BASE COUNT

ORIGIN

Query Match 2.5%; Score 61; DB 125; Length 324;

Best Local Similarity 66.3%; Pred. No. 3.10e-31;

Matches 118; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Db 22 gaagcgnangcaagaataagataatctctttaccacgaagcaaatgagccctggaga 81

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 367 GAAGAAACAGCGCAGAAATCGGACAACTTCGCTCTTCAGCAGTTCGAGG 426

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 82 aagsgtttgagagaccattatccagatgtgttgctgagagaagactagcagccaaaa 141

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 427 CAGTCTTGGCCCAACACTACCTACCCAGATGTCTTCCACGAGAAGAGTACCCATGAAAA 486

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 142 tagatctacctgaagcaagaatacagatgtgtttcttaacgagggccaaatggag 199

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 487 TAAACCTCAGCAAGACCGAGAGTGCAGGTGGTTGTTCCAGACCGAGAGCGGAG 544

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5

LOCUS T09185 386 bp mRNA EST 03-AUG-1993

DEFINITION EST07078 Homo sapiens cDNA clone HIBBQ26 5' end similar to

Oculorhombin (aniridia).

ACCESSION T09185

NID g390213

KEYWORDS EST.

SOURCE

Human clone=HIBBQ26 library=Infant brain, Bento Soares primer=W13

Reverse.

ORGANISM Homo sapiens

Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;


```

/organism="Mus musculus"
/strain="C57BL/6J"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGGAAATTTTTTTTTTTTTTTT
T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
State Univ., from 2 ]; double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT7T3 vector. Library went through one round of
normalization, and was constructed by Bento Soares and
M.Fatima Bonaldo."
/clone="464692"
/clone_lib="Soares mouse embryo NbME13.5 14.5"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"
<1..>572
BASE COUNT      122 a 189 c 164 g  97 t
ORIGIN

Query Match      2.3%; Score 55; DB 3; Length 572;
Best Local Similarity 66.3%; Pred. No. 1.92e-25;
Matches 112; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Db 110 aagcagcgcggaatgaaccacattcaacagcagcagcgcgctgagcgtgta 169
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 372 AACACGGCAGGAATCGGACAACTTCGCTCTTCACAGATTGGAAGCTCTGGAGGCGATC 431

Db 170 tttagcgcacacactaccctgacgccttgtgctgaagagctagctgcgcgtgtaac 229
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 432 TTGTCCCAACACACTACCAGATGCTTCACACAGAGAGAGCTAGCCATGAAATAAAC 491

Db 230 ctcaagtgaagcagctgcaagctctggttcacagacgcgcgcgcgcgcgcgcgcgc 278
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 492 CTCACAGAAGCCAGAGTGCAGGTTTGGTTTCCAGAGAGAGAGCCCAAGT 540

RESULT 8
LOCUS      W89306      604 bp      mRNA      EST      12-SEP-1996
DEFINITION      mf81d10.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
clone 420691 5' similar to gb:X52875 Mouse homeobox gene S8 mRNA
(MOUSE);.
ACCESSION      W89306
VERSION      g1542173
KEYWORDS      EST.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Mus.
REFERENCE      1 (bases 1 to 604)
AUTHORS      Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)

TITLE      Contact: Marra M/Mouse EST Project
JOURNAL      WashU-HHMI Mouse EST Project
COMMENT      Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:255243

```

```

Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 460.
Location/Qualifiers
1..604
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGGAAATTTTTTTTTTTTTTTT
T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
State Univ., from 2 ]; double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT7T3 vector. Library went through one round of
normalization, and was constructed by Bento Soares and
M.Fatima Bonaldo."
/clone="420691"
/clone_lib="Soares mouse embryo NbME13.5 14.5"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"
<1..>604
BASE COUNT      130 a 200 c 171 g 102 t  1 others
ORIGIN

Query Match      2.3%; Score 55; DB 182; Length 604;
Best Local Similarity 66.3%; Pred. No. 1.92e-25;
Matches 112; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Db 113 aagcagcgcggaatgaaccacattcaacagcagcagcgcgctgagcgtgta 172
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 372 AACACGGCAGGAATCGGACAACTTCGCTCTTCACAGATTGGAAGCTCTGGAGGCGATC 431

Db 173 tttagcgcacacactaccctgacgccttgtgctgaagagctagctgcgcgtgtaac 232
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 432 TTGTCCCAACACACTACCAGATGCTTCACAGAGAGAGCTAGCCATGAAATAAAC 491

Db 233 ctcaagtgaagcagctgcaagctctggttcacagacgcgcgcgcgcgcgcgcgcgc 281
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 492 CTCACAGAAGCCAGAGTGCAGGTTTGGTTTCCAGAGAGAGAGCCCAAGT 540

RESULT 9
LOCUS      W17990      424 bp      mRNA      EST      10-SEP-1996
DEFINITION      mb82e05.r1 Soares mouse p3NM19.5 Mus musculus cDNA clone 335936 5'
similar to gb:003873 Mus musculus homeobox (MOUSE);.
ACCESSION      W17990
VERSION      g1292374
KEYWORDS      EST.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Mus.
REFERENCE      1 (bases 1 to 424)
AUTHORS      Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)

TITLE      Contact: Marra M/Mouse EST Project
JOURNAL      WashU-HHMI Mouse EST Project
COMMENT      Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

```


dbb 200 agcgagatccgaaccactctcacacagacgactgcaggagtggagaagatcttc 259
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 376 AGCGAGAATCGGACAACCTTCGTCTTCAAGACTGGAGCTCGAGGCAGTCTTG 435

dbb 260 atttaccactaccocatgtccatactcgaaccaactnnaqccaqntcaacctnc 319

BASE COUNT	55 a	84 c	92 g	78 t
mRNA				
	/iab_nostc- DR10B			
	<1..>309			

Query Match 1.9%; Score 45; DB 180; Length 309;
Best Local Similarity 64.7%; Pred. No. 3.39e-16;
Matches 110; Conservative 0; Mismatches 59; Indels

W P S R L A

(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by Intelligent, Inc.

MPSrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Thu May 29 11:34:49 1997; Maspar time 13.15 Seconds
abular output not generated. 570.251 Million cell updates/sec

Title: >US-08-701-278-2
Description: (1-263) from US08701278.pep
Perfect Score: 1857
Sequence: 1 MFVHCPPQLEGTAPGNHS.....EGSQDKPSPTKEQSEGEKSV 263

Scoring table: PAM 150
Gap 11
Searched: 89912 seqs, 28507787 residues
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: p1r50
1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3
8:unann4 9:unann5 10:unann6 11:unann7 12:unann8
13:unann9 14:unann10 15:unenc 16:unrev
Statistics: Mean 44.651; Variance 107.204; scale 0.417

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	ID	Description
1	399	21.5	A46403	transcription factor
2	392	21.1	I48713	Phox2 homeodomain pr
3	375	20.2	I50413	paired-related homeo
4	374	20.1	I53118	MHOX - mouse
5	374	20.1	I48902	homeobox protein - m
6	374	20.1	I52076	DNA-binding protein
7	373	20.1	I51226	homeodomain protein
8	373	20.1	I27842	homeotic protein smo
9	371	20.0	I48185	gene alx3 protein -
10	369	19.9	I45452	transcription factor
11	370	19.9	I50115	transcription factor
12	369	19.9	I51031	pax-3 protein - mous
13	366	19.7	I51031	homeotic protein S8
14	366	19.7	I48410	gene S8 protein - mo
15	366	19.7	I50413	Prx-2 protein - chic
16	364	19.6	I50250	mab-18 protein (tran
17	364	19.6	I50251	mab-18 protein (tran
18	364	19.6	I50252	vab-3 protein - Caen
19	363	19.5	I47523	cartilage homeoprote
20	360	19.4	I49265	Pax7 - mouse (fragme
21	354	19.1	I51234	gene Pax-6 protein -

22	354	19.1	269 14	S36166	Pax-6 protein - rat	6.90e-37
23	354	19.1	416 13	S37689	Pax-QNR protein - qu	6.90e-37
24	354	19.1	422 13	A56674	paired box homeotic	6.90e-37
25	354	19.1	436 14	S42234	pax-6 protein - mous	6.90e-37
26	354	19.1	447 13	A41644	homeotic protein ani	6.90e-37
27	353	19.0	437 12	S18558	pax protein - zebra	9.97e-37
28	353	19.0	442 12	A57374	homeotic protein PAX	9.97e-37
29	353	19.0	613 12	A26062	segmentation protein	9.97e-37
30	351	18.9	361 14	I49594	homeobox protein - m	2.08e-36
31	348	18.7	80 12	A26332	homeotic protein BSH	6.23e-36
32	347	18.7	449 12	B43698	BSH4 protein - fruit	8.99e-36
33	346	18.6	184 12	S20258	homeotic protein ceh	1.30e-35
34	342	18.4	245 7	A47539	gastrulation-express	5.60e-35
35	342	18.4	288 12	I51620	homeobox protein - A	5.60e-35
36	341	18.4	838 12	I45557	eyeless, long form -	8.07e-35
37	339	18.3	289 14	S35346	otx2 protein - mouse	1.67e-34
38	338	18.2	252 7	A54677	homeotic protein goo	2.41e-34
39	338	18.2	256 7	A42768	gastrulation-express	2.41e-34
40	338	18.2	355 14	I56547	homeodomain 159..341	2.41e-34
41	338	18.2	427 12	A43698	BSH9 protein - fruit	2.41e-34
42	336	18.1	323 16	I50509	orthodenticle-relate	5.00e-34
43	335	18.0	243 12	I51424	homeobox protein - A	7.20e-34
44	334	18.0	243 7	B42768	homeotic protein goo	1.04e-33
45	333	17.9	612 12	A54282	reversed polarity pr	1.49e-33

ALIGNMENTS

RESULT 1
ENTRY A46403 #type complete
TITLE transcription factor with prd-type homeo domain and
ORGANISM Pro/Gln-rich domain-al - Drosophila
DATE 21-Sep-1993 #sequence_revision 21-Sep-1993 #text_change 19-Nov-1993
ACCESSION A46403
REFERENCE A46403
#authors Schneitz, K.; Spielmann, P.; Noll, M.
#journal Genes Dev. (1993) 7:114-129
#title Molecular genetics of aristaless, a prd-type homeo box gene involved in the morphogenesis of proximal and distal pattern elements in a subset of appendages in Drosophila.
#cross-references MUID:93138380
#accession A46403
#molecule_type preliminary
#status nucleic acid
#residues 1-384 #label SCH
#cross-references NCBI:123442; NCBI:123444
#note sequence extracted from NCBI backbone
SUMMARY #length 384 #molecular-weight 41058 #checksum 2627

```

#accession A46403
#status preliminary
#molecule_type nucleic acid
#residues 1-384 #label SCH
#cross-references NCBI:123442: NCBI:123444
#note sequence extracted from NCBI backbone
SUMMARY #length 384 #molecular-weight 41058 #checksum 2627

Query Match 21.5%; Score 399; DB:16; Length 384;
Best Local Similarity 71.8%; Pred. No. 4.05e-44;
Matches 51; Conservative 9; Mismatched 11; Indels 0; Gaps 0;

Db 82 krkqrryrttftsfqleekafsrthypdvftreelamkigtleariqvfwfqrakwr 141
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 30 RRKQRRRTTFALQLEALEAVFAQTHYPDVFTREELAMKINLTARVQVWFQNRRAKWR 89.
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Db 142 kqekvgppqshp 152
|:|:|:|
QY 90 KTERGASDQEP 100

RESULT 2
ENTRY I48713 #type complete
TITLE Phox2 homeodomain protein - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 02-Jul-1996
ACCESSION I48713
REFERENCE I48314
#authors Valarche, I.; Tissler-Seta, J.P.; Hirsch, M.R.; Martinez, S.;

```

Goridis, C.; Brunet, J.F.
Development (1993) 119:881-896
The mouse homeodomain protein Phox2 regulates Ncam promoter activity in concert with Cux/CDP and is a putative determinant of neurotransmitter phenotype.

#cross-references MUID:94244481

#accession I48713

#status preliminary; translated from GB/EMBL/DBJ

#molecule_type mRNA

#residues 1-280 #label RES

#cross-references EMBL:X75014; NID:g402641; CDS_PID:g402642

GENETICS

#note

SUMMARY

gene name Phox2

#length 280 #molecular-weight 29417 #checksum 3267

Query Match 21.1%; Score 392; DB 14; Length 280;

Best Local Similarity 69.9%; Pred. No. 5.48e-43;

Matches 51; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

87 krkqrrrttfsaqkqlervfaethydpdiyreelalkldltearvqvfwqnrakfr 146

30 RRKQRRRTTFAQQLEALEAVFAQTHYDPVFTREELAMKINLTARVQVWFQNRRAKWR 89

Db 147 kceeraasakgaag 159

Qy 90 KTERGASDQEPGA 102

RESULT 3

ENTRY I50413 #type complete

TITLE paired-related homeotic gene product - chicken

ORGANISM #formal_name Gallus gallus #common_name chicken

DATE 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change

ACCESSIONS 150413

REFERENCE 150413

#authors Nohno, T.; Koyama, E.; Myokai, F.; Taniguchi, S.; Ohuchi, H.;

Saito, T.; Noji, S.

#journal Dev. Biol. (1993) 158:254-264

#title The chicken homeobox gene related to Drosophila paired is

#cross-references MUID:93321789

#accession 150413

#status preliminary; translated from GB/EMBL/DBJ

#molecule_type mRNA

#residues 1-244 #label NOH

#cross-references GB:D13433; NID:g222850; CDS_PID:g222851

GENETICS

#note

SUMMARY

gene name Prx-1

#length 244 #molecular-weight 27130 #checksum 5105

Query Match 20.2%; Score 375; DB 13; Length 244;

Best Local Similarity 55.3%; Pred. No. 3.00e-40;

Matches 52; Conservative 16; Mismatches 25; Indels 1; Gaps 1;

91 krkqrrrttfsaqkqlervferthypdafvrdlarrvnltearvqvfwqnrakfr 150

30 RRKQRRRTTFAQQLEALEAVFAQTHYDPVFTREELAMKINLTARVQVWFQNRRAKWR 89

Db 151 rnermlasknasllksysgdvaveqpvprap 184

Qy 90 KTERGASDQEPGA-KEPMAEVTTPPVNRINSPPP 122

RESULT 4

ENTRY I53118 #type complete

TITLE MHox - mouse

ORGANISM #formal_name Mus musculus #common_name house mouse

DATE 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change

ACCESSIONS I53118

REFERENCE I53118

#authors Cserjesi, P.; Lilly, B.; Bryson, L.J.; Wang, Y.; Sassoon,

D.A.; Olson, E.N.

#journal Development (1992) 115:1087-1101

#title MHox: a mesodermally restricted homeodomain protein that binds

#cross-references MUID:93083424

#accession I53118

#status preliminary; translated from GB/EMBL/DBJ

#molecule_type mRNA

#residues 1-217 #label RES

#cross-references GB:L06502; NID:g199583; CDS_PID:g199584

#cross-references MUID:93083424

#length 217 #molecular-weight 24369 #checksum 410

SUMMARY

Query Match 20.1%; Score 374; DB 14; Length 217;

Best Local Similarity 54.7%; Pred. No. 4.35e-40;

Matches 52; Conservative 18; Mismatches 23; Indels 2; Gaps 2;

91 krkqrrrttfsaqkqlervferthypdafvrdlarrvnltearvqvfwqnrakfr 150

30 RRKQRRRTTFAQQLEALEAVFAQTHYDPVFTREELAMKINLTARVQVWFQNRRAKWR 89

Db 151 rnermlasknasllksysgdvaveqpvprap 185

Qy 90 KTERGA-SQEPGA-KEPMAEVTTPPVNRINSPPP 122

RESULT 5

ENTRY I48902 #type complete

TITLE homeobox protein - mouse

ORGANISM #formal_name Mus musculus #common_name house mouse

DATE 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change

ACCESSIONS I48902

REFERENCE I48902

#authors Kern, M.J.; Argao, E.A.; Birkenmeier, E.H.; Rowe, L.B.;

Potter, S.S.

#journal Genomics (1994) 19:334-340

#title Genomic organization and chromosome localization of the

#cross-references MUID:94245205

#accession I48902

#status preliminary; translated from GB/EMBL/DBJ

#molecule_type mRNA

#residues 1-217 #label RES

#cross-references EMBL:U03873; NID:g460124; CDS_PID:g460125

GENETICS

#note

SUMMARY

gene name Pmx

#length 217 #molecular-weight 24369 #checksum 410

Query Match 20.1%; Score 374; DB 14; Length 217;

Best Local Similarity 54.7%; Pred. No. 4.35e-40;

Matches 52; Conservative 18; Mismatches 23; Indels 2; Gaps 2;

91 krkqrrrttfsaqkqlervferthypdafvrdlarrvnltearvqvfwqnrakfr 150

30 RRKQRRRTTFAQQLEALEAVFAQTHYDPVFTREELAMKINLTARVQVWFQNRRAKWR 89

Db 151 rnermlasknasllksysgdvaveqpvprap 185

Qy 90 KTERGA-SQEPGA-KEPMAEVTTPPVNRINSPPP 122

RESULT 6

ENTRY S26076 #type complete

TITLE DNA-binding protein - mouse

ORGANISM #formal_name Mus musculus #common_name house mouse

DATE 25-Feb-1994 #sequence_revision 01-Sep-1995 #text_change

ACCESSIONS S26076

REFERENCE S26076

#authors Kern, M.J.; Witte, D.P.; Valerius, M.T.; Aronow, B.J.;

Potter, S.S.

#journal Nucleic Acids Res. (1992) 20:5189-5195

#title A novel murine homeobox gene isolated by a tissue specific

RESULT	12
ENTRY	S15031 #type complete
TITLE	pax-3 protein - mouse
ORGANISM	#formal_name Mus musculus #common_name house mouse
DATE	21-Nov-1993 #sequence_revision 21-Jul-1995 #text_change 08-Dec-1995
ACCESSIONS	S15031
REFERENCE	S15031
#authors	Goulding, M.D.; Chalepakis, G.; Deutsch, U.; Erselius, J.R.; Gruss, P.
#journal	EMBO J. (1991) 10:1135-1147
#title	Pax-3, a novel murine DNA binding protein expressed during early neurogenesis.
#cross-references	MUID:91216108
#accession	S15031
##status	preliminary
##molecule_type	mRNA
##residues	1-479 #label GOU
CLASSIFICATION	#superfamily paired box homology; homeobox homology
FEATURE	
34-159	#domain paired box homology #label PBH\
220-276	#domain homeobox homology #label HOX
SUMMARY	#length 479 #molecular-weight 52984 #checksum 6402
Query Match	19.9%; Score 369; DB 14; Length 479;
Best Local Similarity	66.7%; Pred. No. 2.76e-39;
Matches	46; Conservative 13; Mismatches 9; Indels 1; Gaps 1;
Db	215 lktqrrsttfaegleeleraferthpydlytreelaqrakltcarqvfvfnsrarrw 274 : ::: : : : : :
QY	29 LRRQRNRTTFALQLEALEAVFAQTHPDTFREELANKINLTERVQVWFQNRRAKW 88 : : : : : :
Db	275 rk-qagaaq 282 : :
QY	89 RKTERGASD 97 : :
RESULT	13
ENTRY	S18038 #type fragment
TITLE	homeotic protein S8 - mouse
ORGANISM	#formal_name Mus musculus #common_name house mouse
DATE	31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 30-Sep-1993
ACCESSIONS	S18038
REFERENCE	S18038
#authors	Opstelten, D.J.E.; Vogels, R.; Robert, B.; Kalkhoven, E.; Zwartkuis, F.; de Laaf, L.; Destree, O.H.; Deschamps, J.; Lawson, K.A.; Meijlink, F.
#submission	submitted to the EMBL Data Library, May 1990
#description	The mouse homeobox gene, S8, is expressed during embryogenesis predominantly in mesenchyme.
#accession	S18038
##molecule_type	mRNA
##residues	1-164 #label OPS
##cross-references	EMBL:X52875
CLASSIFICATION	#superfamily homeobox homology
KEYWORDS	DNA binding; homeobox
FEATURE	
16-72	#domain homeobox homology #label HOX
SUMMARY	#length 164 #checksum 4472
Query Match	19.7%; Score 366; DB 14; Length 164;
Best Local Similarity	70.8%; Pred. No. 8.34e-39;
Matches	46; Conservative 9; Mismatches 10; Indels 0; Gaps 0;
Db	12 kkkqrrnttfssqialeryverthydpdafreelarvnlsearvcvfqnrrakfr 71 .: : : : : : : :
QY	30 RRQRNRTTFALQLEALEAVFAQTHYDPVTREELANKINLTERVQVWFQNRRAKWR 89 .: : : : : : : :
Db	72 rnera 76 : :
QY	90 KTERG 94 : :

Search completed: Thu May 29 11:35:54 1997
Job time : 65 secs.

```
RESULT 14
ENTRY I48410 #type fragment
TITLE gene S8 protein - mouse (fragment)
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
02-Jul-1996
ACCESSIONS I48410
REFERENCE I48410
#authors Opstelten, D.J.; Vogels, R.; Robert, B.; Kalkhoven, E.;
Zwartkruis, F.; de Laaf, L.; Destree, O.H.; Deschamps, J.;
Lawson, K.A.; Weijlink, F.
#journal Mech. Dev. (1991) 34:29-41
#title The mouse homeobox gene, S8, is expressed during
embryogenesis predominantly in mesenchyme.
#cross-references MUID:92001534
#accession I48410
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-164 #label RES
#cross-references EMBL:X52875; NID:g51365; CDS_PID:g51366
JENETICS
#note gene name S8
SUMMARY #length 164 #checksum 4472

Query Match 19.7%; Score 366; DB 14; Length 164;
Best Local Similarity 70.8%; Pred. No. 8.34e-39;
Matches 46; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

Db 12 kkkqrntttfssqlqalervferthypdafvreelarrvnlsearvqvfgnrrakfr 71
Qy 30 RKQRNRTTALQOLEALEAVFAQTHYPDPVFTREELAMKINLTARVQVWFQNRRAKWR 89

Db 72 rnera 76
Qy 90 KTERG 94

RESULT 15
ENTRY S49440 #type complete
TITLE Prx-2 protein chicken
ORGANISM #formal_name gallus gallus #common_name chicken
DATE 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change
07-Jul-1995
ACCESSIONS S49440
REFERENCE S49440
#authors Brouwer, A.; Weijlink, F.
#submission submitted to the EMBL Data Library, June 1994
#description Expression of members of the Prx-family of homeobox genes (S8
and Mhox/K2/Pmx/Phox1) in mouse and chicken embryos.
#accession S49440
#status preliminary
#molecule_type mRNA
#residues 1-165 #label BRO
#cross-references EMBL:X79695
CLASSIFICATION #superfamily homeobox homology
FEATURE
17-73 #domain homeobox homology #label HOX
SUMMARY #length 165 #molecular-weight 18691 #checksum 7828

Query Match 19.7%; Score 366; DB 13; Length 165;
Best Local Similarity 70.8%; Pred. No. 8.34e-39;
Matches 46; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

Db 13 kkkqrntttfssqlqalervferthypdafvreelarrvnlsearvqvfgnrrakfr 72
Qy 30 RKQRNRTTALQOLEALEAVFAQTHYPDPVFTREELAMKINLTARVQVWFQNRRAKWR 89

Db 73 rnera 77
Qy 90 KTERG 94
```

[M][A][S][E][R][V] (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (C) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by Intelligenetics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu May 29 11:33:55 1997; MasPar time 9.24 Seconds
603.427 Million cell updates/sec

abular output not generated.

Title: >US-08-701-278-2
Description: (1-263) from US08701278.pep
Perfect Score: 1857
Sequence: 1 MFYFHCPPQLGTAPEGNHS.....EGSQDKPSPPTKEQGEKSV 263

Scoring table: PAM 150
Gap 11

Searched: 59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot34

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11

Statistics: Mean 46.324; Variance 87.195; scale 0.531

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

sult No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	375	20.2	244	5	HPRI_CHICK	HOMEBOX PROTEIN PRX-	2.18e-51
2	374	20.1	199	7	PMX1_HUMAN	PAIRED MESODERM HOME	3.49e-51
3	374	20.1	217	7	PMX1_MOUSE	PAIRED MESODERM HOME	3.49e-51
4	370	19.9	467	7	PAX7_HUMAN	PAIRED BOX PROTEIN PA	2.27e-50
5	369	19.9	479	7	PAX3_MOUSE	PAIRED BOX PROTEIN PA	3.62e-50
6	369	19.9	479	7	PAX3_HUMAN	PAIRED BOX PROTEIN PA	3.62e-50
7	360	19.4	290	7	PAX7_MOUSE	PAIRED BOX PROTEIN PA	2.42e-48
8	357	19.2	419	4	GSC_DRONE	HOMEBOX PROTEIN GOOS	9.79e-48
9	354	19.1	216	7	PAX6_CHICK	PAIRED BOX PROTEIN PA	3.95e-47
10	354	19.1	416	7	PAX6_COTJA	PAIRED BOX PROTEIN PA	3.95e-47
11	354	19.1	422	7	PAX6_MOUSE	PAIRED BOX PROTEIN PA	3.95e-47
12	354	19.1	422	7	PAX6_HUMAN	PAIRED BOX PROTEIN PA	3.95e-47
13	353	19.0	437	7	PMX6_BRARE	PAIRED BOX PROTEIN PA	6.29e-47
14	353	19.0	613	5	HMFR_DRONE	SEGMENTATION PROTEIN	6.29e-47
15	351	18.9	449	4	GSPB_DRONE	GOOSEBERRY PROXIMAL P	1.59e-46
16	346	18.6	252	10	UNCA_CAEEL	HOMEBOX PROTEIN UNC-	1.62e-45
17	342	18.4	245	4	GSC_CHICK	HOMEBOX PROTEIN GOOS	1.03e-44
18	339	18.3	289	7	OTX2_MOUSE	HOMEBOX PROTEIN OTX2	4.13e-44
19	338	18.2	256	4	GSC_MOUSE	HOMEBOX PROTEIN GOOS	6.56e-44
20	338	18.2	289	7	OTX2_HUMAN	HOMEBOX PROTEIN OTX2	6.56e-44
21	338	18.2	427	4	GSPB_DRONE	GOOSEBERRY DISTAL PRO	6.56e-44
22	337	18.1	240	4	GSC_BRARE	HOMEBOX PROTEIN GOOS	1.04e-43

ID	HPRI_CHICK	STANDARD;	PRT;	244 AA.
AC	Q05437;			
DT	01-JUN-1994 (REL. 29, CREATED)			
DT	01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)			
DT	01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)			
DE	HOMEBOX PROTEIN PRX-1.			
GN	PRX-1			
OS	GALLUS GALLUS (CHICKEN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;			
OC	GALLIFORMES.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-LIMB BUD.			
RX	MEDLINE; 93321789.			
RA	NOHNO T., KOYAMA E., MYOKAI F., TANIGUCHI S., OHUCHI H., SAITO T.,			
RA	NOJI S.;			
RL	DEV. BIOL. 158:254-264(1993).			
CC	-!- FUNCTION: MAY BE INVOLVED IN THE DIFFERENTIATION OF BONE, MUSCLE,			
CC	AND OTHER TISSUES OF MESODERMAL ORIGIN DURING LIMB DEVELOPMENT.			
CC	-!- SUBCELLULAR LOCATION: NUCLEAR.			
CC	CELLS OF THE LIMB BUD, VISCERAL ARCHES AND CRANIOFACIAL PROCESS,			
CC	AND AT LOWER LEVELS, IN CRANIAL MESENCHYME, UPPER AND LOWER			
CC	EYELEIDS, SOMITES AND CARTILAGE OF VERTEBRA.			
CC	-!- SIMILARITY: STRONG WITH OTHER "PAIRED-TYPE" HOMEBOX PROTEINS.			
DR	EMBL; D13433; G222851; -			
DR	HSSP; P02836; IENH.			
DR	PROSITE; PS00027; HOMEBOX.			
KW	HOMEBOX; DNA-BINDING; DEVELOPMENTAL PROTEIN; NUCLEAR PROTEIN.			
FT	DNA BIND 94 153			
SQ	SEQUENCE 244 AA; 27130 MW; 6BD4C3A2 CRC32;			
Query Match 20.2%; Score 375; DB 5; Length 244;				
Best Local Similarity 55.3%; Pred. No. 2.18e-51;				
Matches 52; Conservative 16; Mismatches 25; Indels 1; Gaps 1;				
Db	91 krkqrntttfnsqqlarvferthypdafvredlarrvnltearvqvfnrakfr 150			
Qy	: : : : : : : : : :			
Qy	30 RRKQRNRTTALQLEALEAVFAQTHYPDVTREELAMKINLTREARVQVFNRAKWR 89			
Db	151 rneimlasknasllksygdvtaveqivprpap 184			
Qy	: : : : : : : : : : : :			
Qy	90 KTERGASDQEPGA-KEPMAEVTTPPVNRINSPPP 122			

ALIGNMENTS

[illegible]

DR TRANSFAC; T00396; --
 KW TRANSCRIPTION REGULATION; HOMEOBOX; DNA-BINDING; NUCLEAR PROTEIN;
 RW DEVELOPMENTAL PROTEIN; PAIRED BOX.
 FT DOMAIN 34 161 PAIRED BOX.
 FT DNA_BIND 217 276 HOMEOBOX.
 FT DOMAIN 340 346 POLY-ALA.
 FT CONFLICT 151 152 MISSING (IN REF. 2).
 SQ SEQUENCE 467 AA; 51019 MW; D2C7BCBD CRC32;

Query Match 19.9%; Score 370; DB 7; Length 467;
 Best Local Similarity 66.7%; Pred. No. 2,27e-50;

Matches 46; Conservative 13; Mismatches 9; Indels 1; Gaps 1;

Db 213 lkrqrrrttfaeqllealekaferthypdytreelaqrkltearvqvfnsrrarw 272

QY 29 LRRKQRRNRTTFALQQLALEAVFAQTHYDFVTFREELAMKINLTARVQVWFQNRRAK 88

Db 273 rk-qagaaq 280

QY 89 RKTERGASD 97

RESULT 5

ID PAX3_MOUSE STANDARD; PRT; 479 AA.

AC P24610;

DT 01-MAR-1992 (REL. 21, CREATED)

DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)

DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)

DE PAIRED BOX PROTEIN PAX-3.

GN PAX3 OR PAX-3.

OS MUS MUSCULUS (MOUSE).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;

OC EUTHERIA; RODENTIA.

[1]

RN SEQUENCE FROM N.A.

RX MEDLINE: 91216108.

RA GOULDING M.D.; CHALSPAKIS G., DEUTSCH U., ERSELIUS J.R., GRUSS P.;

RL EMBL J. 10:1135-1147(1991).

[2] SEQUENCE OF 219-260 FROM N.A., AND VARIANT SPLITCH.

RN MEDLINE: 92034997.

RA EPSTEIN D.J.; VEKEMANS M., GROS P.;

RL CELL 67:767-774(1991).

[3]

RN VARIANT SPLITCH.

RX MEDLINE: 94010930.

RA VOGAN K.J.; EPSSTEIN D.J.; TRASLER D.G., GROS P.;

RL GENOMICS 17:364-369(1993).

-1- FUNCTION: PROBABLE TRANSCRIPTION FACTOR ASSOCIATED WITH

DEVELOPMENT OF ALVEOLAR RHABDOMYOSARCOMA.

-1- DEVELOPMENTAL STAGE: EXPRESSED DURING EARLY NEUROGENESIS.

-1- SUBCELLULAR LOCATION: NUCLEAR.

-1- SUBUNIT: CAN BIND AS A HETERODIMER WITH PAX7.

-1- DISEASE: THE SPLITCH (SP) MOUSE MUTANT DISPLAYS DEFECTS IN NEURAL

TUBE CLOSURE IN THE FORM OF EXENCEPHALY AND SPINA BIFIDA.

-1- SIMILARITY: STRONG WITH OTHER "PAIRED-TYPE" HOMEOBOX PROTEINS.

-1- SIMILARITY: CONTAINS A PAIRED BOX DOMAIN.

CC EMBL: X59358; G53592; -.

CC EMBL: S66429; G239201; -.

CC EMBL: S66433; G239203; -.

CC PIR: S15031; S15031.

CC HSSP: P02836; 1ENH.

CC PROSITE: PS00027; HOMEOBOX.

CC PROSITE: PS00034; PAIRED_BOX.

DR TRANSFAC; T00680; -.

KW TRANSCRIPTION REGULATION; PAIRED BOX; DEVELOPMENTAL PROTEIN;

KW NUCLEAR PROTEIN; DNA-BINDING; HOMEOBOX; DISEASE MUTATION.

FT DOMAIN 34 159 PAIRED BOX.

FT DNA_BIND 219 278 HOMEOBOX.

FT VARIANT 42 42 G -> R (IN SPLITCH).

SQ SEQUENCE 479 AA; 52984 MW; 985984B6 CRC32;

Query Match 19.9%; Score 369; DB 7; Length 479;

Best Local Similarity 66.7%; Pred. No. 3,62e-50;
 Matches 46; Conservative 13; Mismatches 9; Indels 1; Gaps 1;

Db 215 lkrqrrrttfaeqllealekaferthypdytreelaqrkltearvqvfnsrrarw 274

QY 29 LRRKQRRNRTTFALQQLALEAVFAQTHYDFVTFREELAMKINLTARVQVWFQNRRAK 88

Db 275 rk-qagaaq 282

QY 89 RKTERGASD 97

RESULT 6

ID PAX3_HUMAN STANDARD; PRT; 479 AA.

AC P23760;

DT 01-NOV-1991 (REL. 20, CREATED)

DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)

DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)

DE PAIRED BOX PROTEIN PAX-3 (HUP2).

GN PAX3.

OS HOMO SAPIENS (HUMAN).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;

OC EUTHERIA; PRIMATES.

[1]

RN SEQUENCE OF 1-29 AND 197-479 FROM N.A.

RA RIETHMAN H.C.; MACINA R.A.;

RL SUBMITTED (JUL-1994) TO EMBL/GENBANK/DBJ DATA BANKS.

[2]

RN SEQUENCE OF 30-195 FROM N.A.

RX MEDLINE: 89305521.

RA BURRI M.; TROMVOUKIS Y., BOPP D., FRIGERIO G., NOLL M.;

RL EMBL J. 8:1183-1190(1989).

[3]

RN SEQUENCE OF 196-392 FROM N.A.

RX MEDLINE: 95072569.

RA TASSABEHJI M.; NEWTON V., LEVERTON K., TURNBULL K., SEEMANOVA E.;

RA KUNZE J., SPERLING K., STRACHAN T., READ A.;

RL HUM. MOL. GENET. 3:1069-1074(1994).

[4]

RN SEQUENCE FROM N.A. (PAX3A AND PAX3B).

RX MEDLINE: 94171226.

RA TSUKAMOTO K.; NAKAMURA Y., NIIKAWA N.;

RL HUM. GENET. 93:270-274(1994).

[5]

RN VARIANT WS1 ALA-63--ILE-67 DEL.

RX MEDLINE: 92168113.

RA TASSABEHJI M.; READ A.P., NEWTON V.E., HARRIS R., BALLING R.;

RA GRUSS P.; STRACHAN T.;

RL NATURE 355:635-636(1992).

[6]

RN VARIANT WS1 LEU-50.

RX MEDLINE: 92168114.

RA BALDWIN C.T.; HOTH C.F., AMOS J.A., DA-SILVA E.O., MILUNSKY A.;

RL NATURE 355:637-638(1992).

[7]

RN VARIANT WS2 ALA-81.

RX MEDLINE: 93258399.

RA TASSABEHJI M.; READ A.P., NEWTON V.E., PATTON M., GRUSS P., HARRIS R.;

RA STRACHAN T.;

RL NAT. GENET. 3:26-30(1993).

[8]

RN VARIANTS WS3 HIS-47 AND WS1 LEU-56.

RX MEDLINE: 93190976.

RA HOTH C.F., MILUNSKY A., LIPSKY N., SHEFFER R., CLARREN S.K.;

RA BALDWIN C.T.;

RL AM. J. HUM. GENET. 52:455-462(1993).

[9]

RN VARIANTS WS1 LEU-45 AND ASP-99.

RX MEDLINE: 95072569.

RA TASSABEHJI M.; NEWTON V.E., LEVERTON K., TURNBULL K., SEEMANOVA E.;

RA KUNZE J., SPERLING K., STRACHAN T., READ A.P.;

RL HUM. MOL. GENET. 3:1069-1074(1994).

[10]

RN

RP VARIANT WS1 VAL-62.
RX MEDLINE: 95135456.
RA PIERPONT J.W., DOOLAN L.D., AMANN K., SNEAD G.R., ERICKSON R.P.;
RL HUM. MUTAT. 4:227-228(1994).
[11]
RN RP VARIANTS WS1 PHE-265 AND GLY-271.
RX MEDLINE: 95126143.
RA LALWANI A.K., BRISTER J.R., FEX J., GRUNDFAST K.M., PLOPLIS B.,
RA SAN AGUSTIN T.B., WILCOX E.R.;
RL AM. J. HUM. GENET. 56:75-83(1995).
[12]
RN RP VARIANT WS3 PHE-84.
RX MEDLINE: 95243235.
RA ZLOFGORA J., LERER I., BAR-DAVID S., ERGAZ Z., ABELIOVICH D.;
RL AM. J. HUM. GENET. 56:1173-1178(1995).
[13]
RN RP VARIANTS WS1 MET-60; GLU-85 AND SER-238.
RX MEDLINE: 96042708.
RA BALDWIN C.T., HOTH C.F., MACINA R.A., MILUNSKY A.;
RL AM. J. MED. GENET. 58:115-122(1995).
[14]
RN RP VARIANT CDHS LYS-47.
RX MEDLINE: 96263735.
RA ASHER J.H. JR., SOMMER A., MORELL R., FRIEDMAN T.B.;
RL HUM. MUTAT. 7:30-35(1996).
[15]
RN RP VARIANT LYS-315.
RA HOL F.A., GEURDS M.P.A., CHATKUP T.S., SHUGART Y.Y., BALLING R.,
RA SCHRANDE-STUMPEL C.T.R.M., JOHNSON W.G., HAMEL B.C.J.,
RA MARIMAN E.C.M.;
RL J. MED. GENET. 33:655-660(1996).
CC -1- FUNCTION: PROBABLE TRANSCRIPTION FACTOR ASSOCIATED WITH
DEVELOPMENT OF ALVEOLAR RHABDYOSARCOMA.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SUBUNIT: CAN BIND AS A HETERODIMER WITH PAX7.
CC -1- ALTERNATIVE PRODUCTS: IN ADDITION TO THE NORMAL PAX3 PROTEIN, THE
GENE CAN PRODUCE, BY ALTERNATIVE SPLICING, TWO TRUNCATED FORMS
WHICH ARE CALLED PAX3A AND PAX3B.
CC -1- DISEASE: DEFECT IN PAX3 ARE THE CAUSE OF WAARDENBURG'S SYNDROME
(WS), AN AUTOSOMAL DOMINANT DISORDER, CHARACTERIZED WITH NON
PROGRESSIVE SENSORY-NEURAL DEAFNESS, AS WELL AS PIGMENTARY CHANGES
OF THE IRIDES AND OF THE HAIR AND SKIN; EACH OF THESE FEATURES
MAY BE UNI- OR BILATERAL. ON THE BASIS OF THE PRESENCE OR ABSENCE
OF DYSTOPIA CANTHORUM (LATERAL DISPLACEMENT OF THE INNER CORNER OF
THE EYE), TYPE I (WS1) AND TYPE II (WS2): DOES NOT MAP TO PAX3) ARE
DISTINGUISHED. TYPE III (WS3 OR KLEIN-WAARDENBURG SYNDROME) IS
CHARACTERIZED WITH WS1 AND UPPER-LIMB INVOLVEMENT INCLUDING MUSCLE
HYPOPLASIA WITH JOINT CONTRACTURES.
CC -1- DISEASE: CRANIOFACIAL-DEAFNESS-HAND SYNDROME (CDHS), IS THOUGHT
TO BE AN AUTOSOMAL DOMINANT DISEASE WHICH COMPRISES ABSENCE OR
HYPOPLASIA OF THE NASAL BONES, HYPOPLASTIC MAXILLA, SMALL AND
SHORT NOSE WITH THIN NARES, LIMITED MOVEMENT OF THE WRIST, SHORT
PALMARAL FISSURES, ULNAR DEVIATION OF THE FINGERS, HYPERTELORISM
AND PROFOUND SENSORY-NEURAL DEAFNESS.
CC -1- DISEASE: RHABDYOSARCOMA-2 (RMS2) IS CHARACTERIZED BY A
CHROMOSOMAL TRANSLOCATION T(2;13)(Q35;Q14) WHICH INVOLVES PAX3 AND
FHR.
CC -1- SIMILARITY: STRONG WITH OTHER "PAIRED-TYPE" HOMEBOX PROTEINS.
DR EMBL: U12263; G555818;
DR EMBL: U12259; G555819;
DR EMBL: U12258; G555819; JOINED.
DR EMBL: U12260; G555819; JOINED.
DR EMBL: U12262; G555819; JOINED.
DR EMBL: X15043; E224055;
DR EMBL: X15252; E224055; JOINED.
DR EMBL: X15253; E224055; JOINED.
DR EMBL: Z29972; -; NOT_ANNOTATED_CDS.
DR EMBL: Z29973; -; NOT_ANNOTATED_CDS.
DR EMBL: Z29974; -; NOT_ANNOTATED_CDS.
DR EMBL: S69369; G545845;
DR EMBL: S69370; G545847;
PIR: S06960; S06960.

DR MIN: 193500; -;
DR MIN: 148820; -;
DR MIN: 268220; -;
DR MIN: 122880; -;
DR PROSITE; PS00027; HOMEBOX.
DR PROSITE; PS00034; PAIRED_BOX.
DR TRANSFAC; T00397; -;
DR TRANSFAC; T00679; -;
KW NUCLEAR PROTEIN; DISEASE MUTATION; CHROMOSOMAL TRANSLOCATION;
KW PROTO-ONCOGENE; ALTERNATIVE SPLICING; POLYMORPHISM.
FT DOMAIN 34 159
FT DNA_BIND 219 278
FT VARSPLIC 197 215
FT VARSPLIC 216 479
FT VARSPLIC 197 206
FT VARSPLIC 207 479
FT VARSPLIC 45 45
FT VARIANT 47 47
FT VARIANT 47 47
FT VARIANT 50 50
FT VARIANT 56 56
FT VARIANT 60 60
FT VARIANT 62 62
FT VARIANT 63 67
FT VARIANT 81 81
FT VARIANT 84 84
FT VARIANT 85 85
FT VARIANT 99 99
FT VARIANT 238 238
FT VARIANT 265 265
FT VARIANT 271 271
FT VARIANT 315 315
SQ SEQUENCE 479 AA; 6EE1491D CRC32;
Query Match 19.9%; Score 369; DB 7; Length 479;
Best Local Similarity 66.7%; Pred. No. 3.62e-50;
Matches 46; Conservative 13; Mismatches 9; Indels 1; Gaps 1;
Db 215 lkrqtrrttftaeqlleeralfetherthypdlytreelaqrakltearvqvfnrtarv 274
Qy 29 LRRQRNRTTFAQLQLEAVEPAQTHYDVFTRTELAMKINLTARVQVFNQRRAKW 88
Db 275 rk-qagang 282
Qy 89 RKTERGASD 97
RESULT 7
ID PAX7 MOUSE STANDARD; PRT; 290 AA.
AC P47239;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE PAIRED BOX PROTEIN PAX-7 (FRAGMENT).
GN PAX7 OR PAX-7.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RX MEDLINE: 91265334.
RA JOSTES B., WALTHER C., GRUSS P.;
RL MECH. DEV. 33:27-37(1990).
CC -1- FUNCTION: PROBABLE TRANSCRIPTION FACTOR. IT MAY HAVE A ROLE IN
MYOGENESIS.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SUBUNIT: CAN BIND AS A HETERODIMER WITH PAX3.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED SPECIFICALLY DURING THE DEVELOPMENT
OF THE NERVOUS AND MUSCULAR SYSTEM.
CC -1- SIMILARITY: STRONG WITH OTHER "PAIRED-TYPE" HOMEBOX PROTEINS.

```

OC GALLIFORMES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94170909.
RA LI H.S., YANG J.M., JACOBSON R.D., PASKO D., SUNDIN O.;
RL DEV. BIOL. 162:181-194(1994).
CC -1- FUNCTION: MAY BE A TRANSCRIPTION FACTOR WITH IMPORTANT FUNCTIONS
CC IN EYE AND NASAL DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: STRONG WITH OTHER "PAIRED-TYPE" HOMEBOX PROTEINS.
CC -1- SIMILARITY: CONTAINS A PAIRED BOX DOMAIN.
DR ENBL; S69508; G545839;
KW TRANSCRIPTION REGULATION; HOMEBOX; DNA-BINDING; NUCLEAR PROTEIN;
KW DEVELOPMENTAL PROTEIN; PAIRED BOX.
FT NON_TER 1
FT DNA_BIND 10 69 HOMEBOX.
SQ SEQUENCE 216 AA; 24469 MW; DAB2288B CRC32;

Query Match 19.18; Score 354; DB 7; Length 216;
Best Local Similarity 57.18; Pred.No. 3.95e-47;
Matches 44; Conservative 18; Mismatches 15; Indels 0; Gaps 0;

Db 6 lkrklqrntsfqgeqlalekeferthydpvdfarerlaakidlpearlqvwsfnrakw 65
QY 29 LRRKQRNRRTFALQLEALEAVFAQTHYPDVTRELAMKINLTEARVQWFQNRRAKW 88
    I:: :|||||:- ::||| | :|||||: || :-|||: ||| |||||
Db 66 rreeklrnqrqsantp 82
    I: I: ::: :||
QY 89 RKTERGASDOEPGAKEP 105

RESULT 10
ID PAX6_COTJJA STANDARD; PRT; 416 AA.
AC P47238;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE PAIRED BOX PROTEIN PAX-6 (PAX-QNR).
GN PAX-6.
OS COTURNIX COTURNUX JAPONICA (JAPANESE QUAIL).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC GALLIFORMES.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-NEURORETINA;
RX MEDLINE; 92366156.
RA MARTIN P., CARRIERE C., DOZIER C., QUATANNENS B., MIRABEL M.A.,
RV VANDENBUNDER B., STEHELIN D., SAULE S.;
RL ONCOGENE 7:1721-1728(1992).
CC -1- FUNCTION: MAY BE A TRANSCRIPTION FACTOR WITH IMPORTANT FUNCTIONS
CC IN EYE AND NASAL DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: STRONG WITH OTHER "PAIRED-TYPE" HOMEBOX PROTEINS.
CC -1- SIMILARITY: CONTAINS A PAIRED BOX DOMAIN.
DR ENBL; X70475; G311772;
KW TRANSCRIPTION REGULATION; HOMEBOX; DNA-BINDING; NUCLEAR PROTEIN;
KW DEVELOPMENTAL PROTEIN; PAIRED BOX.
FT DOMAIN 4 128 PAIRED BOX.
FT DOMAIN 129 209 GLN/GLY-RICH.
FT DNA_BIND 210 269 HOMEBOX.
SQ SEQUENCE 416 AA; 46027 MW; D636633C CRC32;

Query Match 19.18; Score 354; DB 7; Length 416;
Best Local Similarity 57.18; Pred.No. 3.95e-47;
Matches 44; Conservative 18; Mismatches 15; Indels 0; Gaps 0;

Db 206 lkrklqrntsfqgeqlalekeferthydpvdfarerlaakidlpearlqvwsfnrakw 265
QY 29 LRRKQRNRRTFALQLEALEAVFAQTHYPDVTRELAMKINLTEARVQWFQNRRAKW 88
    I:: :|||||:- ::||| | :|||||: || :-|||: ||| |||||
Db 266 rreeklrnqrqsantp 282
    I: I: ::: :||
QY 89 RKTERGASDOEPGAKEP 105

```



```

QY      89 RKTER 93

RESULT 14
ID  HMPR_DROME      STANDARD;          PRT;      613 AA.
AC  P06601;
DT  01-JAN-1988 (REL. 06, CREATED)
DT  01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT  01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE  SEGMENTATION PROTEIN PAIRED.
DN  PRD.
OS  DROSOPHILA MELANOGASTER (FRUIT FLY).
OC  EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE; 87051745.
RA  FRIGERIO G., BURRI M., BOPP D., BAUMGARTNER S., NOLL M.;
RL  CELL 47-735-746(1986).
CC  -1- FUNCTION: PRD IS A PAIR-RULE PROTEIN REQUIRED FOR SEGMENTATION IN
CC  DROSOPHILA. CAPABLE OF SEQUENCE-SPECIFIC DNA-BINDING.
CC  -1- SUBCELLULAR LOCATION: NUCLEAR.
CC  -1- SIMILARITY: STRONG WITH OTHER "PAIRED-TYPE" HOMEOBOX PROTEINS.
CC  -1- SIMILARITY: CONTAINS A PAIRED BOX DOMAIN.
DR  EMBL; M14548; G158160;
DR  FIR; A26062; A26062.
DR  HSSP; P02836; LENH.
DR  FLYBASE; FBGN0003145; PRD.
DR  PROSITE; PS00027; HOMEOBOX.
DR  PROSITE; PS00034; PAIRED_BOX.
DR  TRANSFAC; T00699;
KW  HOMEOBOX; DNA-BINDING; DEVELOPMENTAL PROTEIN; NUCLEAR PROTEIN;
KW  PAIRED BOX; REPEAT; PAIR-RULE PROTEIN; TRANSCRIPTION REGULATION.
FT  DOMAIN 27 151 PAIRED BOX.
FT  DNA_BIND 213 272 HOMEOBOX.
FT  DOMAIN 552 572 10 X 2 AA TANDEM REPEATS OF H-P.
FT  VARIANT 164 164 A -> T.
FT  VARIANT 220 220 F -> I.
SQ  SEQUENCE 613 AA; 65497 MW; 0922346E CRC32;

Query Match 19.0%; Score 353; DB 5; Length 613;
Best Local Similarity 47.1%; Pred. No. 6-29e-47;
Matches 49; Conservative 26; Mismatches 26; Indels 3; Gaps 3;

Db 202 esepgialkrkqrcttfsaqldeleafaertqydpdiyreelaqrntleatrqvfwf 261
: : : : : ||||| ||||| : : : : : ||||| ||||| : : ||||| |||||
QY 23 DFDGDF-LRRKQRNRRTTALQOALEAVFAQTHYPDVFTREELAMKINLTARYQVWF 81
262 snrrarlkrqhtsvsgagpggaasvshvaassslpsvsvssvps 305
||||| : : : : : ||||| ||||| ||||| : : : : : || : : : : :
QY 82 QNRRAKWRKRTERGASDQEPG-AKEPMAEVTTPP-VRNINSPPEG 123

RESULT 15
ID  GSBP_DROME      STANDARD;          PRT;      449 AA.
AC  P09083;
DT  01-NOV-1988 (REL. 09, CREATED)
DT  01-NOV-1988 (REL. 09, LAST SEQUENCE UPDATE)
DT  01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE  GOOSEBERRY PROXIMAL PROTEIN (BSH4).
DN  GSB-P OR GSBA.
OS  DROSOPHILA MELANOGASTER (FRUIT FLY).
OC  EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE; 88112802.
RA  BAUMGARTNER S., BOPP D., BURRI M., NOLL M.;
RL  GENES DEV. 1:1247-1267(1987).
RN  [2]
RP  SEQUENCE OF 12-145 AND 162-241 FROM N.A.
RX  MEDLINE; 87051758.
RA  BOPP D., BURRI M., BAUMGARTNER S., FRIGERIO G., NOLL M.;
RL  CELL 47:1033-1040(1986).
CC  -1- FUNCTION: GOOSEBERRY DISTAL AND PROXIMAL ARE SEGMENT-POLARITY

```